



Objetivo.

Ejercitar las distintas maneras de visualizar datos.

Ejercicios

- 1) Anscombe¹ fue un matemático que hizo hincapié en la importancia de la visualización de datos dado que, a pesar de que las medidas de resumen (media, mediana, etc.) sobre conjuntos de datos distintos pueden ser similares, eso no significa que la manera en que se distribuyen sus datos sean así de similares.

- a. Cargar el dataset de Anscombe. El mismo se puede obtener utilizando la biblioteca **seaborn** de *python*. Para ello, incorporar el siguiente código:

```
import seaborn as sns
anscombe = sns.load_dataset('anscombe')
```

Explorar un poco el dataset ¿Cuántos registros tiene? ¿Qué campos posee?

- b. El dataset posee 4 subconjuntos de datos (dataset=I, II, III y IV). Llamaremos a cada uno de los *subset*. Para cada uno de los subsets obtener sus medidas resumen (media, mediana, cuartiles). ¿Observa diferencias importantes de estas medidas resumen entre los distintos subsets?
 - c. Ahora, para cada uno de los subsets obtener el valor máximo y mínimo que registra cada uno de ellos. Nuevamente, ¿Observa diferencias importantes de estos valores entre los distintos subsets?
 - d. Realizar un histograma de cada subset. Mostrar la *aproximación de la densidad*.
 - e. Realizar un scatterplot de cada subset. ¿Ahora observa diferencias?
 - f. (Opcional) Utilizando la biblioteca *matplotlib* de *python*, generar una única figura, compuesta por cuatro scatterplots, es decir que contenga los datos de los 4 subsets, uno por cada subplot. Utilizar un color diferente para cada subset (hacerlo utilizando estilos distintos para cada uno de ellos).
- 2) Palmer Penguins es un dataset real obtenido de investigaciones realizadas en tres islas diferentes de la Antártida. ¡Felicitaciones! Debido a su entusiasmo por los pingüinos y sus capacidades, ha sido contratado como científico de datos. Su trabajo consiste en:
 - a. averiguar si las islas están pobladas mayormente por alguna especie en particular, o si estas coexisten

¹ Frank Anscombe: https://en.wikipedia.org/wiki/Frank_Anscombe



- b. y en ambos casos deberá notificar en qué proporciones

Es importante mencionar que deberá reportar sus descubrimientos de manera resumida a través de gráficos.

Ahora que ya logró cumplir con su primer objetivo, deberá describir a estos pingüinos a partir de sus características (columnas con las que cuenta en el dataset). Responder justificando su respuesta a través de un gráfico:

- a. ¿Se puede determinar la especie de un pingüino a partir de una sola característica? En caso de que esto sea posible, ¿cuál sería dicha característica? En caso de existir más de una característica que lo permita, comentarlo.
 - b. Dado un pingüino, ¿podría afirmar que el mismo es hembra a partir de sus características (obviamente excluyendo el campo sexo)? En base a su respuesta ¿se puede discriminar a los pingüinos en machos y hembras a partir de UNA característica en particular (excluyendo sexo)?
 - c. Para cada especie en particular de pingüino ¿se puede separar machos de hembras (obviamente excluyendo el campo sexo)?
 - d. Para los pingüinos hembra, a partir de UNA de sus característica en particular (obviamente que no sea a partir del campo especie), ¿se puede predecir a qué especie pertenece? ¿Y combinando más de una característica?
 - e. Ídem punto anterior pero para pingüinos macho.
- 3) Invertir en la bolsa de valores parecería ser un juego de azar. Sin embargo, puede pasar que, teniendo en cuenta ciertos criterios, uno podría ingeniárselas para salir ganando ... o no?. Veamos un ejemplo.
- a. Cargar el dataset de la bolsa de Nueva York (utilizado en clase)
 - b. Elegir una empresa que sea de su interés (cualquiera de las que se encuentran en el dataset).
 - c. Mostrar a través de un gráfico (¿cuál?) la respuesta a la siguiente pregunta: Si hubiese comprado una acción de esa empresa el 24 de diciembre del 2013 a la mañana (ni bien abrió el mercado) y la hubiese vendido ese mismo día cuando el mercado de valores cerró, finalmente ¿hubiese ganado o perdido dinero?
 - d. Responder la misma pregunta del punto anterior pero comprando la acción el 2 de enero al momento del cierre, y vendiendo luego de una semana (al momento de su menor valor de dicho día).
 - e. Digamos que el día de su cumpleaños del año 2007 usted decidió autoregalarse algo, jugado, acciones de alguna empresa que ud eligió, su yo del futuro, digamos en 2011, ¿estuvo feliz, o triste?